

Analyse von multi-temporalen Sentinel-2-Daten zur Klassifizierung von Landbedeckung und Landnutzung am Kaiserstuhl

MARTIN WEINMANN¹, FELIX EHMER¹ & UWE WEIDNER¹

Zusammenfassung: Mit der Verfügbarkeit von zeitlich hochauflösenden Daten wie denen des Sentinel-2-Systems ergibt sich eine Verbesserung der Datenlage hinsichtlich einer multi-temporalen Auswertung durch Klassifizierung. Im Rahmen dieser Arbeit wurden Sentinel-2-Daten anhand des Untersuchungsgebietes des Kaiserstuhls evaluiert. Hierzu wurden zunächst die verfügbaren Sentinel-2-Daten der Jahre 2017 und 2018 hinsichtlich einer Landbedeckungs- und Landnutzungsklassifizierung gesichtet und anschließend erfolgte eine Auswahl von geeigneten Zeitpunkten. Die Daten dieser ausgewählten Zeitpunkte wurden aufbereitet und anschließend klassifiziert, wobei die Wichtigkeit von einzelnen Zeitpunkten und/oder einzelnen Spektralbändern genauer untersucht wurde.

1 Einleitung

Die Sentinel-2-Satelliten werden zur systematischen Erfassung der Erdoberfläche genutzt (SPOTO et al. 2012; KUKAWSKA et al. 2017). Sie bieten im Vergleich zu vorherigen Erdbeobachtungssystemen, die entweder durch eine relativ hohe räumliche Auflösung und eine niedrige Aufnahmezeit (z.B. Landsat-7) oder durch eine niedrige räumliche Auflösung und eine hohe Aufnahmezeit (z.B. MODIS) charakterisiert sind, sowohl hinsichtlich der Auflösung als auch bezüglich der Aufnahmezeit eine verbesserte Verfügbarkeit flächendeckender Multispektraldaten. Dies lässt sich gewinnbringend für die Szenenanalyse anhand von Satellitendaten einsetzen. Während die multi-spektrale Analyse einer Szene mit relativ hoher geometrischer und spektraler Auflösung beispielsweise zur genaueren Kartierung und besseren Unterscheidung von diversen Landbedeckungs- und Landnutzungsklassen genutzt werden kann, erlaubt die multi-temporale Analyse mit hohen Aufnahmezeiten eine genauere Beschreibung von Veränderungen und von dynamischen Prozessen. Durch eine Kombination kann jedoch die multi-temporale Analyse auch zur Verbesserung der mit der multi-spektralen Analyse erzielten Ergebnisse genutzt werden. Dies ist insbesondere dadurch bedingt, dass das spektrale Reflektanzverhalten verschiedener Objekte von zahlreichen Einflussfaktoren abhängt (z.B. vom Wettergeschehen, von der Jahreszeit, von Wachstumszyklen oder von der Oberflächenbeschaffenheit von Objekten) und dadurch systematische Änderungen auftreten können. Durch die Betrachtung des spektralen Reflektanzverhaltens über mehrere Zeitpunkte hinweg werden zuverlässigere Aussagen über die jeweilige Klassenzugehörigkeit möglich.

Im Rahmen dieses Beitrags wird die multi-temporale Analyse von multi-spektralen Sentinel-2-Daten genauer beleuchtet. Dabei wird für das Untersuchungsgebiet des Kaiserstuhls anhand von verfügbaren Sentinel-2-Daten der Jahre 2017 und 2018 eine Analyse hinsichtlich diverser Land-

¹ Karlsruher Institut für Technologie, Institut für Photogrammetrie und Fernerkundung, Englerstraße 7, D-76131 Karlsruhe, E-Mail: [Martin.Weinmann, Uwe.Weidner]@kit.edu, Felix.Ehmer@student.kit.edu

bedeckungs- und Landnutzungsklassen durchgeführt. Nach der Aufbereitung der Rohdaten im Sinne einer atmosphärischen Korrektur sowie einer Analyse der Wolkenbedeckung und korrespondierender Schatten erfolgt eine Merkmalsextraktion, bei welcher anhand der Reflektanzwerte ausgewählter Spektralkanäle sowohl pro Zeitpunkt als auch für die Kombination bezüglich verschiedener Zeitpunkte Merkmalsvektoren abgeleitet werden. Diese Merkmalsvektoren dienen wiederum als Eingabe in einen Random Forest, der für jeden dieser Fälle eine Klassifizierung hinsichtlich diverser Landbedeckungs- und Landnutzungsklassen durchführt. Als Grundlage für das Training des Random Forest und für die Evaluierung der Qualität der erzielten Klassifizierungsergebnisse werden geeignete Trainings- und Kontrollgebiete herangezogen, welche über eine manuelle Annotation gewonnen wurden. Anhand der Trainingsdaten lässt sich auch auf die Wichtigkeit der einzelnen Spektralkanäle hinsichtlich der betrachteten Klassifizierungsaufgabe schließen, wozu im Rahmen dieser Arbeit verschiedene Verfahren verwendet werden. Die für die Klassifizierung erzielten Ergebnisse zeigen das Potential einer multi-temporalen Analyse deutlich auf, wobei für die betrachtete Szene eine signifikante Steigerung der Gesamtgenauigkeit mit bis zu >10% erreicht wurde. Ferner wird deutlich, dass die Qualität der erzielten Klassifizierungsergebnisse durch das Ignorieren der Reflektanzwerte einzelner Spektralkanäle kaum beeinträchtigt wird.

Nach einer kurzen Zusammenfassung des aktuellen Stands der Technik (Kapitel 2) folgt eine detaillierte Erläuterung der verwendeten Methodik zur Analyse von multi-temporalen Sentinel-2-Daten hinsichtlich diverser Landbedeckungs- und Landnutzungsklassen (Kapitel 3). Daraufhin werden die mit dieser Methodik für ein Gebiet um den Kaiserstuhl erzielten Ergebnisse präsentiert (Kapitel 4). Abschließend folgen eine Zusammenfassung sowie ein Ausblick auf zukünftige Arbeiten (Kapitel 5).

2 Stand der Technik

Prinzipiell kann die Klassifizierung von Sentinel-2-Daten hinsichtlich diverser Landbedeckungs- und Landnutzungsklassen in Analogie zu Standardverfahren der Bildauswertung erfolgen. Wurden Sentinel-2-Daten für einen bestimmten Zeitpunkt erfasst, können sie als Bildstapel interpretiert werden und die Reflektanzwerte der einzelnen Spektralkanäle pro Pixel als Einträge von Merkmalsvektoren definiert werden. Ferner können auch spektrale Indizes wie z.B. der Normalized Difference Vegetation Index (NDVI) oder der Normalized Difference Water Index (NDWI) für die Definition von Merkmalsvektoren genutzt werden (VALERO et al. 2016) sowie eine bestimmte Auswahl von relevanten Kanälen (ZHANG et al. 2017; WEINMANN & WEIDNER 2018; WEINMANN & WEIDNER 2019). Die entsprechend definierten Merkmalsvektoren können wiederum als Eingangswerte für diverse Klassifizierungsverfahren genutzt werden, wobei in der Regel Standardklassifizierungsverfahren wie eine Support Vector Machine (CORTES & VAPNIK 1995) oder ein Random Forest (BREIMAN 2001) genutzt werden, welche mittlerweile auch in vielen Softwarepaketen vorhanden sind. Mittlerweile werden aber auch oftmals leistungsfähigere Deep-Learning-Ansätze zur Klassifizierung von Sentinel-2-Daten hinsichtlich diverser Landbedeckungs- und Landnutzungsklassen genutzt (ZHU et al. 2017). Da derartige Verfahren darauf basieren, eine komplexe Optimierungsaufgabe unter Verwendung vieler Millionen Parameter durchzuführen, wird in der Regel eine große Menge an repräsentativen Trainingsdaten benötigt

bis eine angemessene Generalisierungsfähigkeit auf neue Daten erreicht wird. Um eine aufwändige manuelle Annotation zu vermeiden, können beispielsweise auch OpenStreetMap-Daten für das Training einer Standard-CNN-Architektur genutzt werden (LIEBEL 2018).

Um zusätzlich den Einfluss verschiedener Jahreszeiten und den Einfluss von Wachstumszyklen sowie der damit jeweils verbundenen Änderung des spektralen Reflektanzverhaltens vieler Objekttypen (insbesondere bezüglich diverser Vegetationsklassen) bei der Klassifizierung zu berücksichtigen, können beispielsweise für jedes Pixel die Reflektanzwerte der einzelnen Spektralkanäle pro Zeitpunkt zur Definition von Merkmalsvektoren genutzt werden. Die für die kombinierte Betrachtung von Zeitpunkten resultierenden Merkmalsvektoren können dann wiederum mit Standardklassifizierungsverfahren wie z.B. einer Support Vector Machine oder einem Random Forest den jeweiligen Klassen zugeordnet werden (GÓMEZ et al. 2016). Dies lässt sich auch direkt auf die Schätzung der Verteilung von Baumarten in bewaldeten Ökosystemen anhand von multi-temporalen Sentinel-2-Daten übertragen (KARASIAK et al. 2017). Im Gegensatz zu derartigen Standardverfahren lässt sich eine multi-temporale Analyse von Sentinel-2-Daten auch über eine Rekonstruktion der Zeitreihe realisieren, was beispielsweise zur genaueren Analyse von landwirtschaftlichen Nutzflächen genutzt werden kann (BRUZZONE et al. 2017). Hierbei kann mittels eines nicht-parametrischen Regressionsmodells in Form eines Multi-Layer Perceptrons eine Sentinel-2-Zeitreihe in Tagesschritten rekonstruiert und als Eingabe in eine Support Vector Machine gegeben werden. Neben der spektralen und temporalen Information in den Daten können auch räumliche Informationen in den Satellitenbildern mittels Zusammenhangskomponenten einbezogen werden (SOLANO-CORREA et al. 2017), was die Separierung von einzelnen Feldern und deren Analyse mit einer hohen zeitlichen Auflösung ermöglicht. Alternativ kann die Modellierung von dynamischen Prozessen auch über die Nutzung von Hidden Markov Models (HMMs) (SIACHALOU et al. 2015) oder über eine Erweiterung eines Conditional Random Fields (CRFs) durch den zusätzlichen Einbezug temporaler Interaktionen (HOBERG et al. 2015) erfolgen. Ferner ist auch eine Anwendung von modernen Deep-Learning-Ansätzen für die Klassifizierung von multi-temporalen Sentinel-2-Daten möglich. Diesbezüglich bietet sich insbesondere die Nutzung von „Long Short-Term Memory (LSTM)“ Netzwerken (HOCHREITER & SCHMIDHUBER 1997) an, deren wesentliche Idee im Gegensatz zu herkömmlichen rekurrenten neuronalen Netzen auf dem Einbringen eines „langen Kurzzeitgedächtnisses“ basiert. Die Nutzung derartiger Ansätze zur Landbedeckungs- und Landnutzungsklassifizierung ermöglicht eine Berücksichtigung von zeitlichen Charakteristika und führt dadurch im Vergleich zu Verfahren, bei denen lediglich Daten einzelner Zeitpunkte genutzt werden, zu besseren Klassifizierungsergebnissen (RUBWURM & KÖRNER 2017), erfordert jedoch wiederum eine große Menge an annotierten Daten als Grundlage für das Training.

3 Methodik

Als Eingangsdaten erhält die verwendete Methodik Rohdaten in Form von multi-temporalen Sentinel-2-Daten. Zunächst erfolgt dann eine Datenvorverarbeitung, um aus den verfügbaren Rohdaten über eine geeignete Korrektur auf die entsprechenden Reflektanzwerte der einzelnen Spektralkanäle zu schließen (Kapitel 3.1). Auf der Basis dieser Reflektanzwerte werden anschließend Merkmalsvektoren definiert (Kapitel 3.2), welche wiederum als Eingangsgrößen für

eine Klassifizierung dienen (Kapitel 3.3). Als Ergebnis liefert die verwendete Methodik schließlich eine thematische Karte, welche die Klassenzugehörigkeiten einzelner Pixel des betrachteten Satellitenbildes widerspiegelt. In einigen Fällen bietet es sich an, die Klassifizierung lediglich auf der Grundlage einer bestimmten Auswahl an Spektralkanälen durchzuführen und dabei diese Auswahl anhand von automatisierten Verfahren zu treffen (Kapitel 3.4).

3.1 Datenvorverarbeitung

Zunächst werden die Reflektanzwerte, die zu den gegebenen Spektralkanälen korrespondieren und später zur Definition von Merkmalsvektoren verwendet werden, aus den verfügbaren Rohdaten über eine atmosphärische Korrektur mittels „Sen2Cor“ (<https://step.esa.int/main/third-party-plugins-2/sen2cor/>) ermittelt, wobei verschiedene Parameter eingestellt werden können. Im Rahmen dieser Arbeit werden Effekte von Cirrus-Wolken korrigiert, ein digitales Höhenmodell eingebunden und die bidirektionale Reflexionsverteilungsfunktion (BRDF, von engl. „Bidirectional Reflectance Distribution Function“) berücksichtigt. Zudem wird die Bodenauflösung („Ground Sampling Distance“) für alle Bänder auf eine Auflösung von 20 m angepasst. Um bei der Klassifizierung Effekte durch Wolkenbedeckung oder den Schatten von Wolken zu vermeiden, wird abschließend noch über alle verfügbaren Zeitpunkte geprüft, inwiefern für die entsprechenden Datenpunkte mindestens einmal eine derartige Störung vorliegt (EHMER 2019). Das Verschneiden dieser Information über alle betrachteten Zeitpunkte liefert eine Wolken- und Schattenmaske, über die entsprechende Datenpunkte für alle nachfolgenden Datenverarbeitungsschritte verworfen werden können.

3.2 Merkmalsextraktion

Basierend auf den abgeleiteten Reflektanzwerten erfolgt pro Zeitpunkt eine Auswahl von bestimmten Spektralkanälen für die weiteren Betrachtungen (WEINMANN et al. 2018). Hierbei werden nur die Reflektanzwerte ausgewählt, welche zu den Spektralkanälen B2 (zentrale Wellenlänge: 490 nm), B3 (560 nm), B4 (665 nm), B5 (705 nm), B6 (740 nm), B7 (783 nm), B8a (865 nm), B11 (1610 nm) und B12 (2190 nm) korrespondieren. Reflektanzwerte, welche zu den Spektralkanälen B1 (443 nm), B9 (945 nm) und B10 (1375 nm) korrespondieren, werden nicht berücksichtigt, da sie Bereiche im elektromagnetischen Spektrum abdecken, in denen die atmosphärische Durchlässigkeit nur gering ist oder der Wasserdampf das Verhalten dominiert. Der Spektralkanal B8 (842 nm) wird nicht berücksichtigt, da er mit dem Spektralkanal B8a überlappt, aber deutlich breiter und daher weniger aussagekräftig ist (siehe Abb. 1). Die Reflektanzwerte der ausgewählten Spektralkanäle werden weiterhin zur Definition von Merkmalsvektoren genutzt, wobei eine Konkatenation dieser Reflektanzwerte für jeden Zeitpunkt erfolgt. Liegen entsprechende Reflektanzwerte für mehrere Zeitpunkte vor, erfolgt eine Konkatenation unter Berücksichtigung aller Zeitpunkte.

3.3 Klassifizierung

Die definierten Merkmalsvektoren dienen wiederum als Eingangswerte für ein Klassifizierungsverfahren, welches es erlaubt, die jeweiligen Klassenzugehörigkeiten zu ermitteln. Im Rahmen

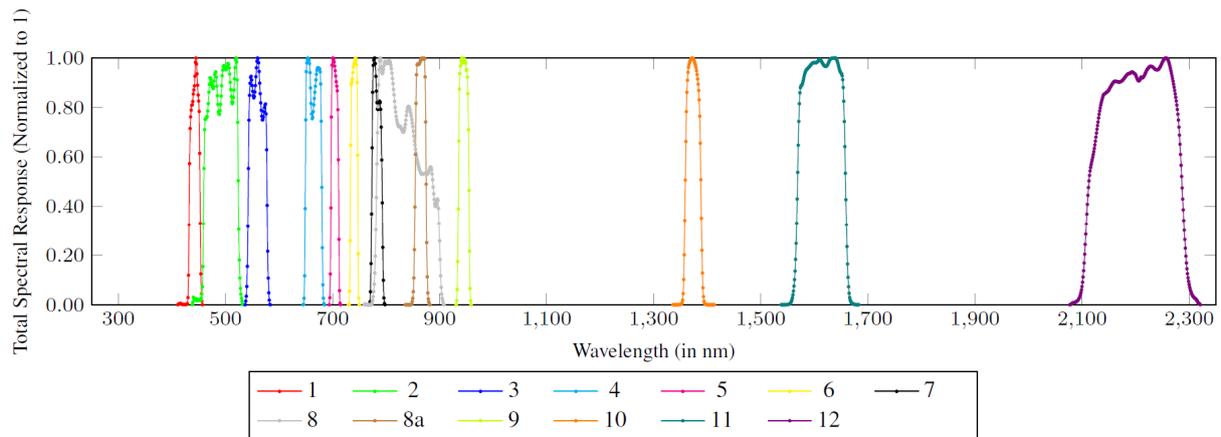


Abb. 1: Eigenschaften der einzelnen Spektralkanäle bei multi-spektralen Sentinel-2-Daten (Weinmann et al. 2018): die Darstellung zeigt die sog. „Sentinel-2 Spectral Response Functions (S2-SRFs)“

dieser Arbeit wird für die Klassifizierung ein überwachtes Lernverfahren in Form eines Random Forest (BREIMAN 2001) verwendet.

In der Trainingsphase wird anhand von Trainingsbeispielen unter Verwendung eines Bootstrap Aggregating („Bagging“) ein Ensemble von randomisierten Entscheidungsbäumen erzeugt, d.h. jeder Entscheidungsbaum wird auf einer zufällig anderen Untermenge der Trainingsdaten trainiert. Beim Training jedes Entscheidungsbaumes erfolgt eine sukzessive Aufteilung der jeweils betrachteten Trainingsdaten in immer kleinere Untermengen, so dass jede resultierende Untermenge „so rein wie möglich“ ist. Diese Aufteilung wird in Form einer Sequenz von Entscheidungen formalisiert.

In der Testphase, d.h. bei der Betrachtung der zu klassifizierenden Merkmalsvektoren, wird dann bei jedem Entscheidungsbaum die entsprechende, durch das Training festgelegte Sequenz von Entscheidungen durchlaufen, bis ein Blatt des Entscheidungsbaumes mit einer entsprechenden Hypothese bezüglich einer Klassenzugehörigkeit erreicht wird. Eine Betrachtung der Hypothesen („Votes“) über alle Entscheidungsbäume führt zu Klassenwahrscheinlichkeiten und der Mehrheitsentscheid („Majority Vote“) liefert schließlich eine robuste Hypothese für eine geeignete Zuweisung einer Klassenzugehörigkeit.

3.4 Selektion von relevanten Spektralkanälen

Um als Grundlage für die Klassifizierung eine sinnvolle Auswahl von Spektralkanälen zu treffen, können verschiedene Strategien der Merkmalsselektion herangezogen werden (GUYON & ELISSEFF 2003; SAEYS et al. 2007; ZHAO et al. 2010; WEINMANN & WEIDNER 2019). Am Beispiel der Analyse von multi- und hyper-spektralen Daten lassen sich relevante Spektralkanäle durch 1) eine intensive Interaktion mit einem Klassifizierungsverfahren, 2) interne Metriken eines Klassifizierungsverfahrens (welche während des Trainings ermittelt werden) oder 3) unabhängig von Klassifizierungsverfahren durch aufgestellte Relevanzmetriken ableiten (WEINMANN & WEIDNER 2019). Im Rahmen dieser Arbeit wurde aus Effizienzgründen ein Verfahren aus der letzteren Kategorie verwendet, welches auf der kombinierten Betrachtung verschiedener Relevanzmetriken (wie z.B. Distanz-, Informations-, Abhängigkeits- oder Korrelationsmetriken) basiert und somit eine Abschätzung der Wichtigkeit einzelner Spektralkanäle hinsichtlich der be-

trachteten Klassifizierungsaufgabe erlaubt (WEINMANN & WEIDNER 2019). Zudem wird im Random Forest eine interne Metrik herangezogen.

4 Ergebnisse

Im Folgenden werden zunächst das Untersuchungsgebiet und die verwendete Datengrundlage beschrieben (Kapitel 4.1). Anschließend wird die Bestimmung von Trainings- und Kontrollgebieten als Grundlage für das Training von Klassifizierungsverfahren und die Evaluierung von Klassifizierungsergebnissen erläutert (Kapitel 4.2). Daraufhin werden die zur Bewertung der Leistungsfähigkeit der vorgestellten Methodik verwendeten Evaluierungsmetriken beschrieben (Kapitel 4.3). Abschließend erfolgt die Präsentation der erzielten Ergebnisse (Kapitel 4.4).

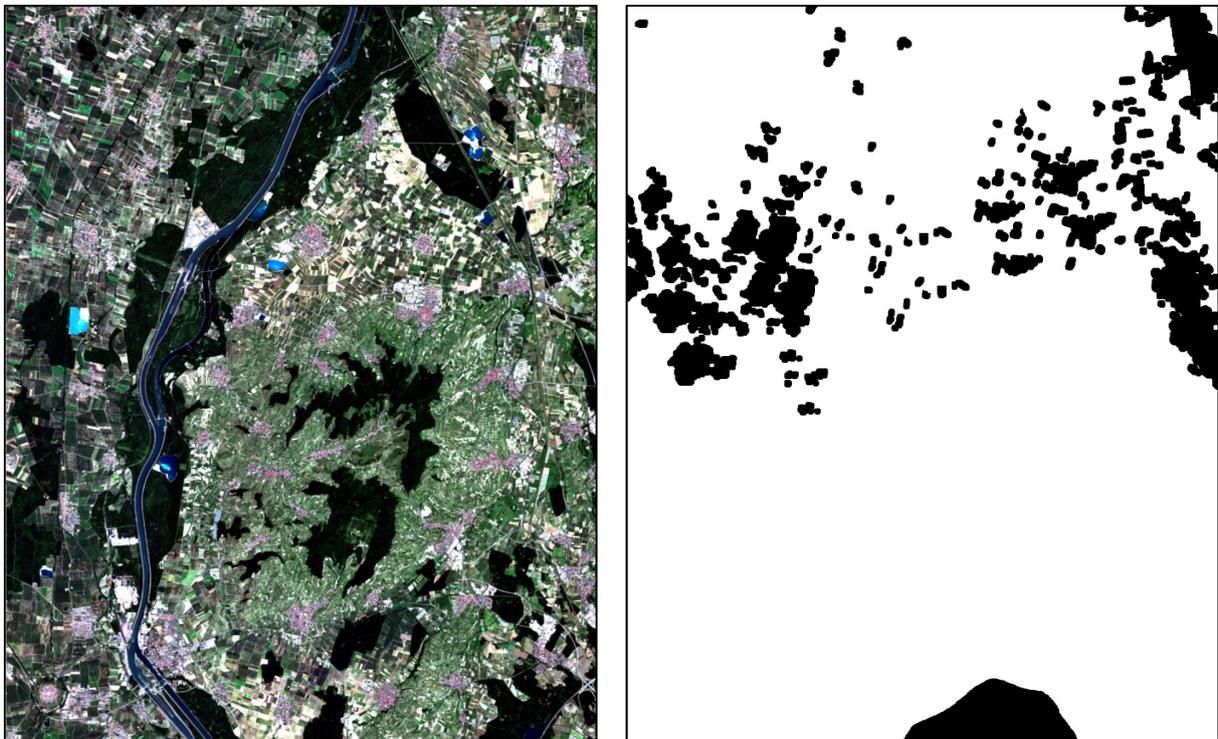


Abb. 2: Falschfarbendarstellung des Untersuchungsgebietes um den Kaiserstuhl für den Zeitpunkt des 19.06.2017 unter Verwendung der Kanalrepräsentation B4-B3-B2 der entsprechenden multispektralen Sentinel-2-Daten (links) sowie die aus 12 verschiedenen Zeitpunkten abgeleitete Wolken- und Schattenmaske (rechts)

4.1 Untersuchungsgebiet

Als Untersuchungsgebiet wurde die Region des Kaiserstuhls gewählt (siehe Abb. 2, links). Der Kaiserstuhl und sein Umland zeichnen sich durch seine vielfältige Vegetation aus. Besonders das Weinanbaugebiet am Kaiserstuhl sowie die hohe landwirtschaftliche Nutzung im Umland prägen diese Region. Aber auch der Rhein westlich des Kaiserstuhls und die vielen Siedlungen haben

ihren Anteil an der facettenreichen Landbedeckung. Insgesamt weist das Untersuchungsgebiet Höhendifferenzen von bis zu knapp 400 m auf.

Für die durchgeführten Untersuchungen wurden Sentinel-2-Daten für 12 verschiedene Zeitpunkte herangezogen, von denen sechs im Jahr 2017 (10.04., 10.05., 19.06., 19.07., 23.08., 22.09.) und sechs im Jahr 2018 (20.04., 05.05., 09.07., 24.07., 03.08., 12.09.) liegen und jeweils die wesentlichen Wachstumszyklen abdecken. Anhand dieser Daten wurde eine Wolken- und Schattenmaske erstellt (siehe Abb. 2, rechts), über die entsprechend fehlerbehaftete Datenpunkte für alle nachfolgenden Datenverarbeitungsschritte verworfen werden.

4.2 Auswahl von Trainings- und Kontrollgebieten

Als Grundlage für das Training eines Klassifizierungsverfahrens und die Evaluierung der Qualität der erzielten Klassifizierungsergebnisse diente eine hinreichende Anzahl von repräsentativen Trainings- und Kontrollgebieten, welche den Klassen *Siedlung*, *Industrie*, *Wasser*, *Wald*, *Wein*, *Wiese* und *Feld* zugeordnet werden können. Diese Trainings- und Kontrollgebiete wurden aus den vorliegenden Satellitenbildern über eine manuelle Annotation von Polygonen mit möglichst homogenen Charakteristika der eingeschlossenen Pixel abgeleitet. Zunächst wurden 29 Trainingsgebiete ausgewählt, welche den definierten Klassen entsprechen. Diese Trainingsgebiete umfassen insgesamt 6298 Pixel und entsprechen damit einer Fläche von 251,92 ha. Anhand dieser Trainingsgebiete mit bekannten Klassenzugehörigkeiten wurde pro Pixel ein Trainingsbeispiel bestehend aus dem entsprechenden Merkmalsvektor und der korrespondierenden Klassenzugehörigkeit definiert. Um eine quantitative Evaluierung der vorgestellten Methodik zu ermöglichen, wurden zudem 46 Kontrollgebiete definiert, welche insgesamt 11514 Pixel und damit eine Fläche von 460,56 ha abdecken. Anhand der Kontrollgebiete mit bekannten Klassenzugehörigkeiten werden wiederum Kontrollbeispiele definiert, die aus dem entsprechenden Merkmalsvektor und der korrespondierenden Klassenzugehörigkeit bestehen. Die entsprechenden Verteilungen über die einzelnen Klassen sind in Tab. 1 und Tab. 2 angegeben.

Tab. 1: Verteilung der Trainingsbeispiele über die definierten Klassen.

Klasse	# Beispiele	Fläche [ha]	Anteil [%]
Siedlung	999	39,96	15,86
Industrie	854	34,16	13,56
Wasser	975	39,00	15,48
Wald	929	37,16	14,75
Wein	898	35,92	14,26
Wiese	604	24,16	9,59
Feld	1039	41,56	16,50
Σ	6298	251,92	100,0

Tab. 2: Verteilung der Kontrollbeispiele über die definierten Klassen.

Klasse	# Beispiele	Fläche [ha]	Anteil [%]
--------	-------------	-------------	------------

Siedlung	1582	63,28	13,74
Industrie	1475	59,00	12,81
Wasser	1305	52,20	11,33
Wald	1936	77,44	16,81
Wein	2226	89,04	19,33
Wiese	1000	40,00	8,69
Feld	1990	79,60	17,28
Σ	11514	460,56	100,0

4.3 Evaluierungsmetriken

Neben der qualitativen Bewertung der erzielten Klassifizierungsergebnisse in Form von thematischen Karten wurde auch eine quantitative Bewertung durchgeführt, bei welcher die Gesamtgenauigkeit (OA) und der κ -Index sowie die klassenweisen Metriken des Recall-Werts und des Precision-Werts herangezogen wurden. Aus den letzteren beiden wurde für jede Klasse der entsprechende F₁-Wert abgeleitet und zur Bewertung über alle Klassen der Mittelwert gebildet (mF₁).

4.4 Ergebnisse

Im Rahmen dieser Arbeit wurden Sentinel-2-Daten, welche für das Gebiet des Kaiserstuhls für verschiedene Zeitpunkte vorliegen, sowohl pro Zeitpunkt als auch für die Kombination bezüglich verschiedener Zeitpunkte analysiert. Dabei erfolgte eine Klassifizierung für die gesamte verfügbare Szene sowie eine quantitative Bewertung der erzielten Ergebnisse anhand der definierten Kontrollgebiete.

Die unter Berücksichtigung einzelner Zeitpunkte erzielten Klassifizierungsergebnisse sind in Tab. 3 zusammengefasst und zeigen starke Variationen bezüglich der betrachteten Evaluierungsmetriken. Der generelle Trend zeigt mit einer OA von 76-84% relativ schlechte Klassifizierungsergebnisse für den ersten Zeitpunkt im April jedes betrachteten Jahres sowie für den letzten Zeitpunkt im September jeden betrachteten Jahres, während bei den Zeitpunkten dazwischen pro Jahr zumeist signifikant bessere Klassifizierungsergebnisse mit einer OA von 88-91% vorliegen. Die jeweils besten Ergebnisse pro Jahr sind in Abb. 3 dargestellt.

Die erzielten Ergebnisse für die Kombination von Sentinel-2-Daten verschiedener Zeitpunkte als Grundlage für die Klassifizierung sind in Abb. 4, Abb. 5 und Tab. 4 angegeben. Die Kombination der Sentinel-2-Daten für die jeweils sechs verschiedenen Zeitpunkte pro Jahr liefert mit einer OA von 93,56% und 94,36% eine signifikante Verbesserung der Klassifizierungsergebnisse. Demgegenüber ist mit einer OA von 94,90% lediglich eine leichte Verbesserung gegeben, wenn die Sentinel-2-Daten für alle 12 verschiedenen Zeitpunkte als Grundlage für die Klassifizierung genutzt werden.

Tab. 3: Klassifizierungsergebnisse unter Berücksichtigung der Sentinel-2-Daten von einzelnen Zeitpunkten in den Jahren 2017 und 2018.

Zeitpunkte	OA [%]	κ [%]	mF ₁ [%]
10.04.2017	82,35	78,65	80,85
10.05.2017	87,68	84,99	85,07
19.06.2017	89,19	86,78	87,81
19.07.2017	80,36	76,01	80,38
23.08.2017	88,00	85,29	84,24
22.09.2017	80,15	75,89	79,05
20.04.2018	76,25	71,51	76,48
05.05.2018	85,39	82,27	83,71
09.07.2018	89,55	87,24	87,49
24.07.2018	87,84	85,12	87,44
03.08.2018	90,97	88,92	89,59
12.09.2018	83,61	80,10	81,86



Abb. 3: Erzielte Klassifizierungsergebnisse unter Berücksichtigung der Sentinel-2-Daten für die Zeitpunkte des 19.06.2017 (links) und des 03.08.2018 (rechts): die Farben entsprechen den Klassen *Siedlung* (in Rot), *Industrie* (in Grau), *Wasser* (in Blau), *Wald* (in Dunkelgrün), *Wein* (in Rosa), *Wiese* (in Hellgrün) und *Feld* (in Gelb). Für die schwarzen Bereiche liegen aufgrund der Verwendung einer Wolken- und Schattenmaske keine Ergebnisse vor.

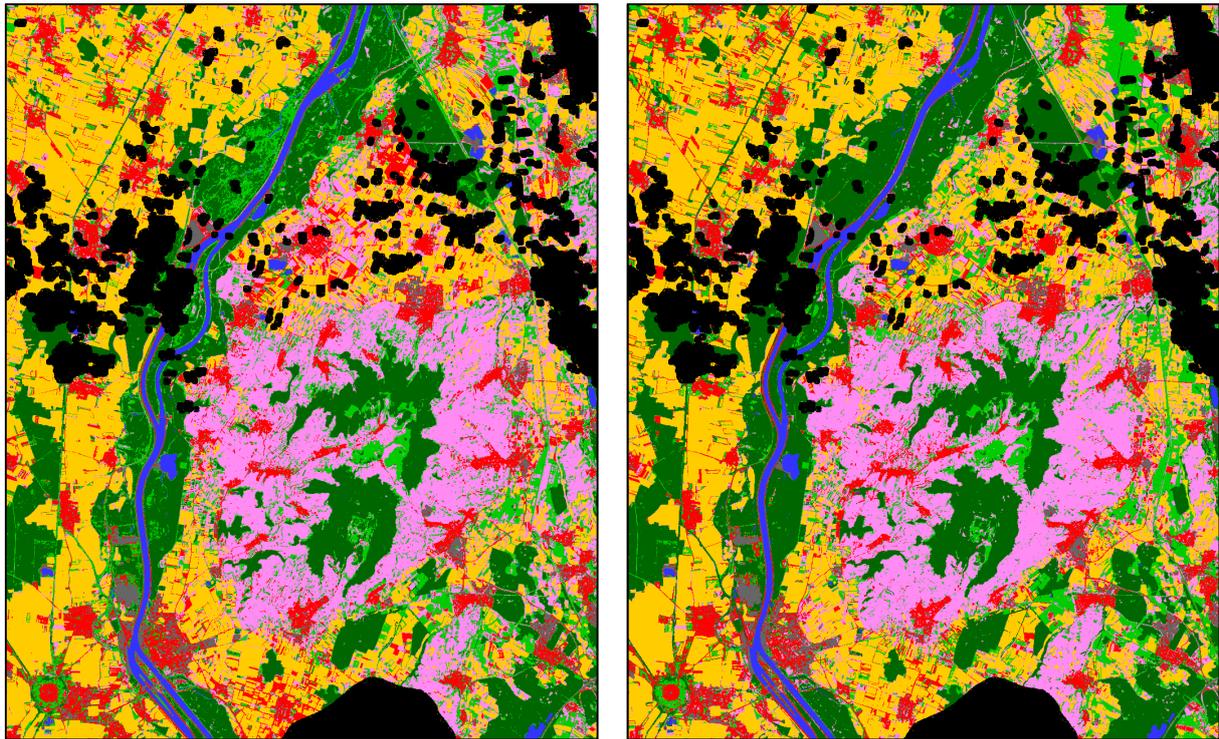


Abb. 4: Klassifizierungsergebnisse für die Jahre 2017 (links) und 2018 (rechts) unter Berücksichtigung der Sentinel-2-Daten für jeweils sechs Zeitpunkte: die Farben entsprechen den Klassen *Siedlung* (in Rot), *Industrie* (in Grau), *Wasser* (in Blau), *Wald* (in Dunkelgrün), *Wein* (in Rosa), *Wiese* (in Hellgrün) und *Feld* (in Gelb). Für die schwarzen Bereiche liegen aufgrund der Verwendung einer Wolken- und Schattenmaske keine Ergebnisse vor

Tab. 4: Klassifizierungsergebnisse unter Berücksichtigung der Sentinel-2-Daten für jeweils sechs verschiedene Zeitpunkte pro Jahr sowie für die Kombination aller Zeitpunkte.

Zeitpunkte	OA [%]	κ [%]	mF ₁ [%]
alle in 2017	93,56	92,11	91,47
alle in 2018	94,36	93,09	92,16
alle in 2017 und 2018	94,90	93,75	92,66

Die erzielten Ergebnisse für die Kombination der Sentinel-2-Daten für jeweils drei verschiedene Zeitpunkte pro Jahr sind in Tab. 5 angegeben. Dabei wird ersichtlich, dass sowohl für das Jahr 2017 als auch für das Jahr 2018 die Berücksichtigung der ersten drei Zeitpunkte zu deutlich besseren Klassifizierungsergebnissen führt als die Berücksichtigung der letzten drei Zeitpunkte. Der Unterschied beträgt hier bezüglich der OA ca. 2-4 %. Insgesamt liegt jedoch eine geringere Qualität der Klassifizierungsergebnisse vor im Vergleich zu den Ergebnissen für die Kombination der Sentinel-2-Daten für die jeweils sechs verschiedenen Zeitpunkte pro Jahr und für die Kombination der Sentinel-2-Daten für alle 12 verfügbaren Zeitpunkte.

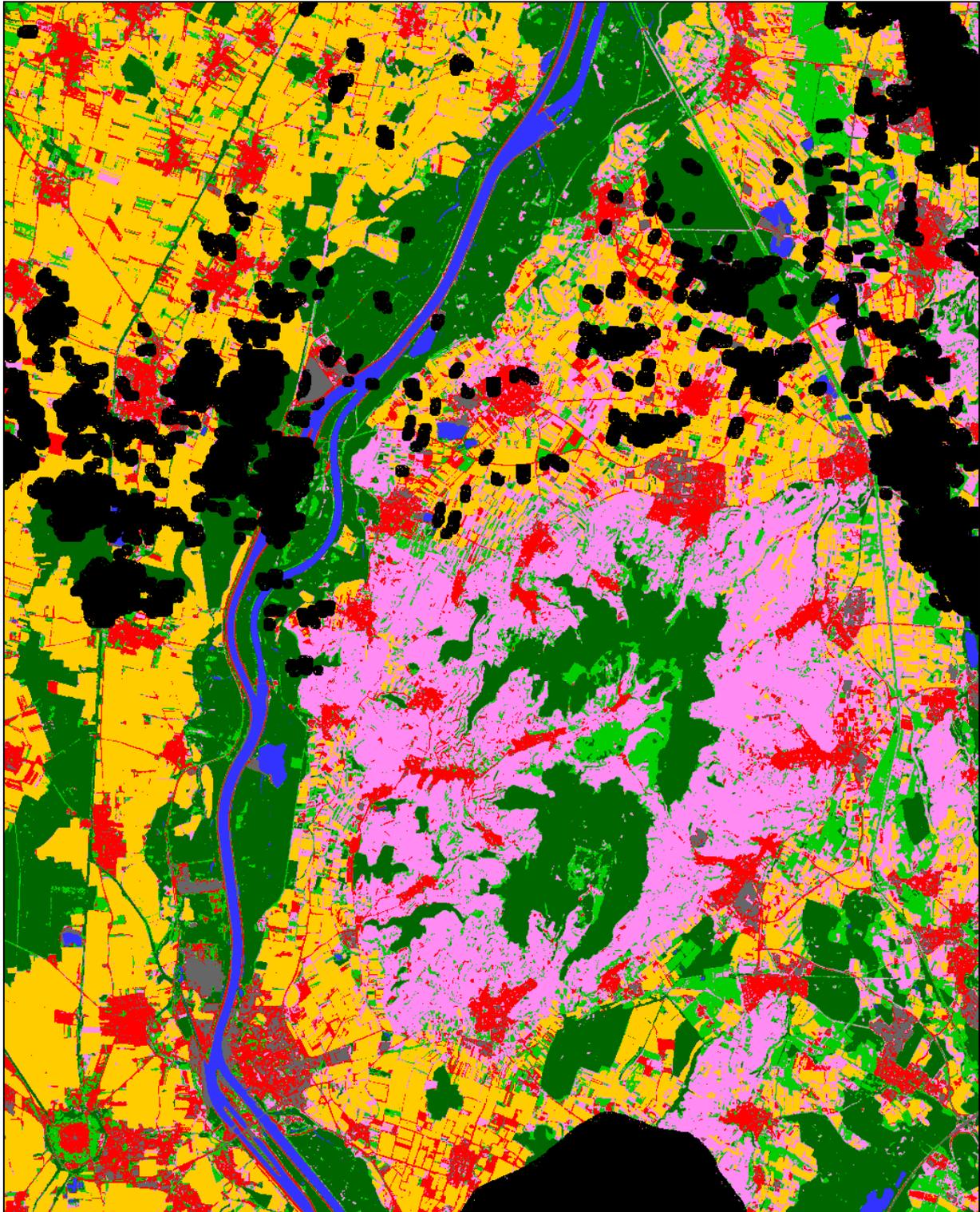


Abb. 5: Klassifizierungsergebnisse unter Berücksichtigung der Sentinel-2-Daten für alle 12 verfügbaren Zeitpunkte: die Farben entsprechen den Klassen *Siedlung* (in Rot), *Industrie* (in Grau), *Wasser* (in Blau), *Wald* (in Dunkelgrün), *Wein* (in Rosa), *Wiese* (in Hellgrün) und *Feld* (in Gelb).

Tab. 5: Klassifizierungsergebnisse unter Berücksichtigung der Sentinel-2-Daten für jeweils drei verschiedene Zeitpunkte pro Jahr.

Zeitpunkte in 2017	OA [%]	κ [%]	mF₁ [%]
10.04., 10.05., 19.06.	92,43	90,75	90,36
19.07., 23.08., 22.09.	88,48	85,96	85,40

Zeitpunkte in 2018	OA [%]	κ [%]	mF₁ [%]
20.04., 05.05., 09.07.	92,55	90,89	90,34
24.07., 03.08., 12.09.	90,44	88,31	88,59

Unter Verwendung der abgeleiteten Relevanzmetrik sowie der internen Metrik des Random Forest scheinen insbesondere die Spektralkanäle B2, B3, B4 und B5 als relevant bezüglich der betrachteten Klassifizierungsaufgabe. Werden daher zunächst die Reflektanzwerte der Spektralkanäle B11 und B12 während der Klassifizierung ignoriert und die Reflektanzwerte der Spektralkanäle B2, B3, B4, B5, B6, B7 und B8a herangezogen, zeigen die Klassifizierungsergebnisse eine etwas geringere Qualität bei einem Verlust von etwa 1-2% bezüglich der OA im Vergleich zu den für die Nutzung aller Spektralkanäle erzielten Ergebnissen (siehe Tab. 6). Werden demgegenüber die Reflektanzwerte der Spektralkanäle B6, B7 und B8a während der Klassifizierung ignoriert und die Reflektanzwerte der Spektralkanäle B2, B3, B4, B5, B11 und B12 herangezogen, beträgt der Verlust bezüglich der OA unter 1% im Vergleich zu den für die Nutzung aller Spektralkanäle erzielten Ergebnissen (siehe Tab. 7). Werden schließlich die Reflektanzwerte der Spektralkanäle B6, B7, B8a, B11 und B12 während der Klassifizierung ignoriert und lediglich die Reflektanzwerte der Spektralkanäle B2, B3, B4 und B5 herangezogen, beträgt der Verlust bezüglich der OA wiederum etwa 1-2% im Vergleich zu den für die Nutzung aller Spektralkanäle erzielten Ergebnissen (siehe Tab. 8).

Tab. 6: Klassifizierungsergebnisse unter Berücksichtigung der Sentinel-2-Daten für jeweils sechs verschiedene Zeitpunkte pro Jahr sowie für die Kombination aller Zeitpunkte: die pro Zeitpunkt verwendeten Spektralkanäle sind durch B2, B3, B4, B5, B6, B7 und B8a gegeben.

Zeitpunkte	OA [%]	κ [%]	mF₁ [%]
alle in 2017	92,26	90,53	89,93
alle in 2018	93,14	91,60	90,91
alle in 2017 und 2018	93,68	92,26	91,35

Tab. 7: Klassifizierungsergebnisse unter Berücksichtigung der Sentinel-2-Daten für jeweils sechs verschiedene Zeitpunkte pro Jahr sowie für die Kombination aller Zeitpunkte: die pro Zeitpunkt verwendeten Spektralkanäle sind durch B2, B3, B4, B5, B11 und B12 gegeben.

Zeitpunkte	OA [%]	κ [%]	mF₁ [%]
alle in 2017	93,43	91,96	91,64
alle in 2018	93,55	92,09	91,82
alle in 2017 und 2018	93,98	92,62	92,47

Tab. 8: Klassifizierungsergebnisse unter Berücksichtigung der Sentinel-2-Daten für jeweils sechs verschiedene Zeitpunkte pro Jahr sowie für die Kombination aller Zeitpunkte: die pro Zeitpunkt verwendeten Spektralkanäle sind durch B2, B3, B4 und B5 gegeben.

Zeitpunkte	OA [%]	κ [%]	mF ₁ [%]
alle in 2017	92,95	91,37	91,21
alle in 2018	92,42	90,71	90,39
alle in 2017 und 2018	92,83	91,23	90,88

5 Zusammenfassung und Ausblick

In diesem Beitrag wurde die Analyse von multi-temporalen Sentinel-2-Daten hinsichtlich diverser Landbedeckungs- und Landnutzungsklassen behandelt. Dabei wurden für das Untersuchungsgebiet des Kaiserstuhls anhand von verfügbaren Sentinel-2-Daten der Jahre 2017 und 2018 über die Reflektanzwerte ausgewählter Spektralkanäle sowohl pro Zeitpunkt als auch für die Kombination bezüglich verschiedener Zeitpunkte Merkmalsvektoren abgeleitet und jeweils als Eingabe für einen Random Forest verwendet. Letzterer lieferte für die resultierenden Fälle eine Klassifizierung hinsichtlich der definierten Landbedeckungs- und Landnutzungs-klassen. Die vergleichenden Betrachtungen der erzielten Klassifizierungsergebnisse zeigten das große Potential einer multi-temporalen Analyse deutlich auf, wobei für das betrachtete Gebiet durch die Kombination der Reflektanzwerte ausgewählter Spektralkanäle bezüglich verschiedener Zeitpunkte im Vergleich zur Betrachtung der Reflektanzwerte ausgewählter Spektralkanäle pro Zeitpunkt eine signifikante Steigerung der Gesamtgenauigkeit mit bis zu >10% erreicht wurde. Ferner konnte gezeigt werden, dass die Qualität der erzielten Klassifizierungsergebnisse insbesondere durch die Information der Reflektanzwerte in den Spektralkanälen B2, B3, B4 und B5 bedingt ist und durch das Ignorieren der Reflektanzwerte anderer Spektralkanäle kaum beeinträchtigt wird.

Im Rahmen zukünftiger Arbeiten ist eine zusätzliche Betrachtung von klassenweisen Evaluierungsmetriken geplant, da dies eine detailliertere Analyse des Verhaltens der erzielten Klassifizierungsergebnisse hinsichtlich einzelner Landbedeckungs- und Landnutzungsklassen ermöglicht. Zudem wird die Erweiterung der Trainingsdaten angestrebt, so dass im Vergleich zu den mit der vorgestellten Methodik für ein Untersuchungsgebiet erzielten Ergebnissen auch die Leistungsfähigkeit von modernen Deep-Learning-Ansätzen analysiert werden kann. Dabei ist auch der Einfluss von fehlerbehafteten Zuordnungen von Klassenzugehörigkeiten bei den zugrundeliegenden Trainingsdaten („Label Noise“) zu untersuchen, da diese in der Regel durch ein effizienteres manuelles Annotieren großer Szenen bedingt sind und oftmals einen nicht vernachlässigbaren Anteil an den Trainingsdaten ausmachen.

6 Literaturverzeichnis

- BREIMAN, L., 2001: Random forests. *Machine Learning* **45**(1), 5-32.
- BRUZZONE, L., BOVOLO, F., PARIS, C., SOLANO-CORREA, Y. T., ZANETTI, M. & FERNÁNDEZ-PRIETO, D., 2017: Analysis of multitemporal Sentinel-2 images in the framework of the ESA Scientific Exploitation of Operational Missions. 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1-4.
- CORTES, C. & VAPNIK, V., 1995: Support-vector networks. *Machine Learning* **20**(3), 273-297.
- EHMER, F., 2019: Untersuchung zur Nutzung von Sentinel2-Daten für die Region des Kaiserstuhls. Bachelorarbeit, Institut für Photogrammetrie und Fernerkundung, Karlsruher Institut für Technologie.
- GÓMEZ, C., WHITE, J. C. & WULDER, M. A., 2016: Optical remotely sensed time series data for land cover classification: a review. *ISPRS Journal of Photogrammetry and Remote Sensing* **116**, 55-72.
- GUYON, I. & ELISSEEFF, A., 2003: An introduction to variable and feature selection. *Journal of Machine Learning Research* **3**, 1157-1182.
- HOBERG, T., ROTTENSTEINER, F., FEITOSA, R. Q. & HEIPKE, C., 2015: Conditional random fields for multitemporal and multiscale classification of optical satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing* **53**(2), 659-673.
- HOCHREITER, S. & SCHMIDHUBER, J., 1997: Long short-term memory. *Neural Computation*, **9** (8), 1735-1780.
- KARASIAK, N., SHEEREN, D., FAUVEL, M., WILLM, J., DEJOUX, J. F. & MONTEIL, C., 2017: Mapping tree species of forests in southwest France using Sentinel-2 image time series. 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1-4.
- KUKAWSKA, E., LEWINSKI, S., KRUPINSKI, M., MALINOWSKI, R., NOWAKOWSKI, A., RYBICKI, M. & KOTARBA, A., 2017: Multitemporal Sentinel-2 data – Remarks and observations. 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1-4.
- LIEBEL, L., 2018: Deep convolutional neural networks for semantic segmentation of multispectral Sentinel-2 satellite imagery: an open data approach to large-scale land use and land cover classification. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V.*, Band **27**, T. P. Kersten, E. Gülch, J. Schiewe, T. H. Kolbe & U. Stilla (Hrsg.), 38. Wissenschaftlich-Technische Jahrestagung der DGPF, 7.-9. März 2018 in München, Tagungsband auf CD-ROM, 790-793.
- RUBWURM, M. & KÖRNER, M., 2017: Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 1496-1504.
- SAEYS, Y., INZA, I. & LARRANAGA, P., 2007: A review of feature selection techniques in bioinformatics. *Bioinformatics* **23**(19), 2507-2517.
- SIACHALOU, S., MALLINIS, G. & TSAKIRI-STRATI, M., 2015: A hidden Markov models approach for crop classification: linking crop phenology to time series of multi-sensor remote sensing data. *Remote Sensing* **7**(4), 3633-3650.

- SOLANO-CORREA, Y. T., BOVOLO, F., BRUZZONE, L. & FERNÁNDEZ-PRIETO, D., 2017: Spatio-temporal evolution of crop fields in Sentinel-2 satellite image time series. 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 1-4.
- SPOTO, F., SY, O., LABERINTI, P., MARTIMORT, P., FERNANDEZ, V., COLIN, O., HOERSCH, B. & MEYGRET, A., 2012: Overview of Sentinel-2. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1707-1710.
- VALERO, S., MORIN, D., INGLADA, J., SEPULCRE, G., ARIAS, M., HAGOLLE, O., DEDIEU, G., BONTEMPS, S., DEFOURNY, P. & KOETZ, B., 2016: Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions. Remote Sensing **8**(1), 55:1-55:21.
- WEINMANN, M., MAIER, P. M., FLORATH, J. & WEIDNER, U., 2018: Investigations on the potential of hyperspectral and Sentinel-2 data for land-cover / land-use classification. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **IV-1**, 155-162.
- WEINMANN, M. & WEIDNER, U., 2018: Erste Erfahrungen mit der Nutzung von Sentinel-2 Daten zur multispektralen und multitemporalen Analyse mit verschiedenen Klassifizierungsverfahren. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **27**, T. P. Kersten, E. Gülch, J. Schiewe, T. H. Kolbe & U. Stilla (Hrsg.), 38. Wissenschaftlich-Technische Jahrestagung der DGPF, 7.-9. März 2018 in München, Tagungsband auf CD-ROM, 58-70.
- WEINMANN, M. & WEIDNER, U., 2019: Relevance assessment of spectral bands for land cover and land use classification: A case study involving multispectral Sentinel-2-like and hyperspectral data. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band **28**, T. P. Kersten (Hrsg.), 39. Wissenschaftlich-Technische Jahrestagung der DGPF, 20.-22. Februar 2019 in Wien, Tagungsband auf CD-ROM, 138-153.
- ZHANG, T., SU, J., LIU, C., CHEN, W. H., LIU, H. & LIU, G., 2017: Band selection in Sentinel-2 satellite for agriculture applications. 23rd International Conference on Automation and Computing (ICAC), 1-6.
- ZHAO, Z., MORSTATTER, F., SHARMA, S., ALELYANI, S., ANAND, A. & LIU, H., 2010: Advancing Feature Selection Research – ASU Feature Selection Repository. Technical Report, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA, 2010.
- ZHU, X. X., TUIA, D., MOU, L., XIA, G., ZHANG, L., XU, F. & FRAUNDORFER, F., 2017: Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine **5**(4), 8-36.